Graph Based Pattern Recognition

Exercise 4

Basis

- Chapter 7

Submission

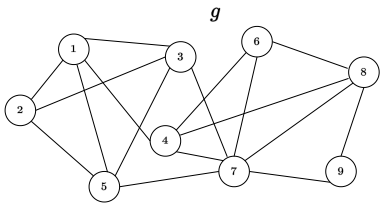
- The submission takes place online on ILIAS.
- Solutions to the theory tasks must be submitted as `*.pdf` file. Other formats will not be accepted.
- Source code for the implementation tasks must be submitted as `*.py` files. Source code that cannot be executed will not be accepted.
- Individual submissions or submissions in teams of two are allowed (hand in only one copy per group). In the source code file, include the *names and matriculation numbers* of both group members in the first two lines as comments.

Dates

- Briefing: 19.04.2023
- Submission: 26.04.2023
- Debriefing: 26.04.2023

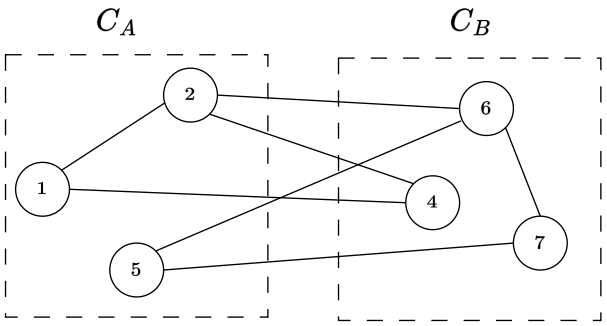
Theoretical Tasks

1. Given the following graph g and three graph clusterings. Compute the graph density, the cut size (e.g., $cut(A, B)$), the intra-cluster density for each partition (e.g., $\delta_{int}(A)$), the general intra-cluster induced by the clustering (e.g., $\delta_{int}(g|A, B)$), and the inter-cluster density (e.g., $\delta_{ext}(g|A, B)$). After computing these metrics, you should analyze the values obtained and make an argument for which of the proposed clusterings is preferred.

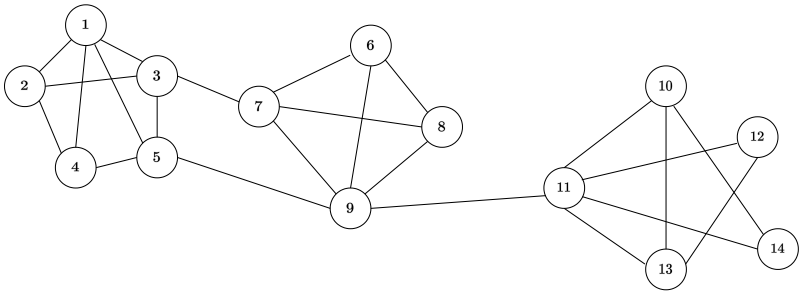


1) $A = \{1, 2, 3, 4\}$ $B = \{5, 6, 7, 8, 9\}$
2) $C = \{1, 2, 3, 5\}$ $D = \{4, 6, 7, 8, 9\}$
3) $E = \{1, 2, 3, 6, 8\}$ $F = \{4, 5, 7, 9\}$

2. Which node pair would be changed in the following bisection in the Kernighan-Lin algorithm? Apply one turn of the Kernighan-Lin algorithm and provide the values E_v, I_v, and D_v for all nodes v in C_A and C_B as well as all values $R_{a,b}$ for all pair of nodes.



3. Compute a spectral graph bisection (see Alg. 10 in the lecture notes) according to the Fiedler vector of the following graph. Apply the spectral bisection algorithm twice, with a threshold of $\theta = 0$, to identify three clusters. First, apply the algorithm to the entire graph to determine the initial partitioning. After obtaining this partitioning, apply the spectral bisection algorithm once again to the larger of the two clusters. For each step report the clustering and the Fiedler vector you obtain.



Implementation Tasks

1. In this exercise series, the goal is to implement Prim's Minimum Spanning Tree (MST) algorithm (see Alg. 12 in the lecture notes). First, go to the ILIAS's webpage of the course and download/unzip `Exercise_4.zip` in your `PR_Lecture` folder. Then, navigate to `PR_lecture/Exercise_4/ex4.py` and complete the missing part of the source code.

   Once you have implemented the MST algorithm apply it on the weighted graph in `PR_lecture/Exercise_4/graphs/graph_00.graphml` and save the spanning tree (i.e., the list of covered edges) as a list of tuples in `PR_lecture/Exercise_4/results/spanning_tree.txt`.

Submission

For the coding part, you must submit a `.zip` file containing the following files. Additionally, include your solution for the theoretical tasks as `*.pdf` in the same `.zip` file.